

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
 - TEXT CUT OFF AT TOP, BOTTOM OR SIDES
 - FADED TEXT
 - ILLEGIBLE TEXT
 - SKEWED/SLANTED IMAGES
 - COLORED PHOTOS
 - BLACK OR VERY BLACK AND WHITE DARK PHOTOS
-
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**



PATENT ABSTRACTS OF JAPAN

(11) Publication number: **10134059 A**(43) Date of publication of application: **22 . 05 . 98**

(51) Int. Cl.

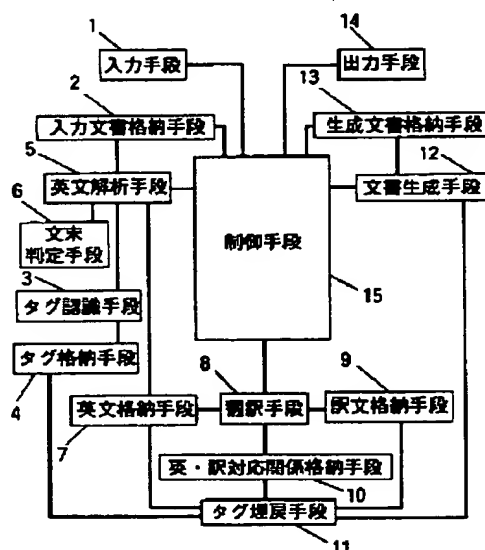
G06F 17/28**G06F 12/00****// G06F 13/00**(21) Application number: **08292331**(22) Date of filing: **05 . 11 . 96**(71) Applicant: **MATSUSHITA ELECTRIC IND CO LTD**(72) Inventor: **AKIYAMA SHIGEKO****(54) HTML DOCUMENT PROCESSOR AND PROCESSING METHOD THEREFOR****(57) Abstract:**

PROBLEM TO BE SOLVED: To view an original text and a translation on the same display screen and to improve user's usability by removing tags in an HTML document and translating sentences to be translated, and generating a new HTML document which has the translation object sentences given the tags back and contains an added translation.

SOLUTION: An English sentence analyzing means 5 analyzes English sentences of document data stored in an input document storage means 2 by using a tag recognition means 3 and a sentence end deciding means 6 and stores the English sentences to be translated in an English sentence storage means 7. Further, the tag recognition means 3 when recognizing a tag stores its tag information in a tag storage means 4. The English sentences stored in the English sentence storage means 7 are translated by a translating means 8 and the obtained translation is stored in a translation storage means 9. A document preparing means 12 embeds the tags in the English sentences stored in the English sentence storage means 7 and the translation stored in the translation storage means 9 according to the

information on the tags stored in the tag storage means 4 by using a tab embedding-back means 11 to generate an HTML document.

COPYRIGHT: (C)1998,JPO



Our Ref: OP1081-US

Prior Art Reference:

Japanese Patent Laid-Open Publication No. 10-134059

Date of Publication: May 22, 1998

Title: HTML DOCUMENT PROCESSING APPARATUS AND
HTML DOCUMENT PROCESSING METHOD

Patent Application No. 8-292331

Filing Date: November 5, 1996

Applicant: ID No. 000005821

MATSUSHITA DENKI SANGYO KABUSHIKI KAISHA
Kadoma-shi, Osakau-fu, Japan

Inventor: Kaoruko AKIYAMA

c/o Matsushita Denki Sangyo Kabushiki Kaisha
Kadoma-shi, Osaka-fu, Japan

Partial Translation

[ABSTRACT]

[Object] To provide a HTML document processing apparatus for analyzing and translating an inputted HTML document, and generating a HTML document containing an original text and a translation thereof and outputting it.

[Solving Means] An apparatus includes an English text analyzing means 5 which recognizes a tag in a HTML document which has been inputted from input means 1, by using tag recognizing means 3, and fetches a text to be translated; tag storing means 4 for storing the recognized tag and a position of the tag; English text storing means 7 for storing the fetched text to be translated; translation means 8 for translating the English text to be translated, which has been stored in the English text storing means, into other language; translated text storing means 9 for storing a translated text resulted from a translation; document generating means 12 which, based on the tag and the tag position stored in the tag storing means, embeds the tag into the

text to be translated, which has been stored in the English text storing means 7, by using tag embedding means 11, and generating a HTML document containing the added translation which has been stored in the translated text storing means 9; and output means 14 for outputting the generated HTML document.

Pertinent Description

[0021]

Fig. 3 is a flow chart of an operation of a HTML document creating apparatus according to an embodiment of the present invention.

[0022]

As shown in Fig. 3, a HTML document is inputted via the input means 1 (Step 1). In this input operation, the input means 1 may be by an FD (floppy disk) having a loading function for loading from a storage medium.

[0023]

The HTML document inputted from the input means 1 is stored in the input document storing means 2 in the form of a text (Step 2).

[0024]

The English text analyzing means 5 performs analysis of the English text with respect to the document data stored in the input document storing means 2 by using the tag recognition means 3 and the text end determining means 6, and stores the English text to be translated in the English text storing means 7. Further, when the tag is recognized by the tag recognition means 3, the recognized tag information is stored in the tag storing means 4 by the English text analyzing means 5 (Step 3).

[0025]

After making the English text analysis in Step 3, the English text stored in the English text storing means 7 is translated by

the translation means 8, and a translated text resulted from the translation is stored in the translated text storing means 9. Further, when the translation means 8 performs the translation, the corresponding relationship between the English text stored in the English text storing means 7 and the translated text stored in the translated text storing means 9 is stored in the English-translation corresponding relationship storing means 10 (Step 4).

[0026]

Thereafter, the document generating means 12, based on the tag information stored in the tag storing means 4, embeds the tag in the English text stored in the English text storing means 7 and in the translated text stored in the translated text storing means 9 by using the tag embedding means 11, thereby to generate a HTML document and stores it in the generated document storing means 13 (Step 5). Details will be described later.

[0027]

When the HTML document containing the translated text is generated, the HTML document stored in the generated document storing means 13 is outputted to the output means 14 (Step 6).

[0028]

The operation of the HTML document processing apparatus is as described above. The operation of each of Step 3 and Step 5 of Fig. 3 will be described in detail hereinbelow.

[0029]

Fig. 4 is a flow chart of the English text analyzing means according to an embodiment of the present invention, and explaining in detail an operation of the English text analysis performed in Step 3 of Fig. 3.

[0030]

In Fig. 4, whether the analysis process of the document data stored in the input document storing means 2 is finished or not is determined (Step 11).

[0031]

When determined in Step 11 that the analysis is finished, the process goes to Yes, and moves to the translation process of the above-described Step 4 of Fig. 3. Further, when determined that the analysis is not finished, the process goes to No, and determines a kind of character by using the tag recognition means 3, namely, determines whether the character is English or tag (Step 12).

[0032]

When determined in Step 12 that the character is English, then whether the character is the text end character or not is determined by using the text end determining means 6 (Step 13).

[0033]

When determined in Step 13 that the character is the text end character, the process goes to Yes, fetches the English text, and stores the fetched English text, as the text to be translated, in the English text storing means 7 (step 14). When determined in Step 13 that the character is not the text end character, the process goes to No and returns to Step 11.

[0034]

On the other hand, when, in Step 12, the kind of character is determined to be the tag, the tag information is stored in the tag storing means 4 (Step 15), and whether the tag is the tag of the text end or not is determined by using the text end determining means 6 (Step 16).

[0035]

When determined in Step 16 that the tag is the text end tag, the process goes to Yes, and moves to Step 14. When determined that it is not the text end tag, the process goes to No, and returns to Step 11.

[0036]

The above described operation of Fig. 4 is performed for all data stored in the input document storing means 2.

[0037]

Fig. 5 is a flow chart of the document generating means according to an embodiment of the present invention, explaining in detail the operation of HTML document generating process in Step 5 of Fig. 3.

[0038]

As shown in Fig. 5, whether the HTML document generating process for the data stored in the English text storing means 7 and in the translated text storing means 9 is finished or not is determined (Step 21). When determined that the generating process is finished, the process moves to the document output process in the above-described Step 6 of Fig. 3.

[0039]

When determined in Step 21 that the generating process is not finished, the process goes to No, and determines whether embedding of the tag is necessary or not, based on the tag and the tag information stored in the tag storing means 4, by using the tag embedding means 11 (Step 22).

[0040]

When determined in Step 22 that embedding of the tag is necessary, the process advances to Yes, and determines a kind of text to be processed (Step 23).

[0041]

When determined in Step 23 that the kind of text is English, the tag is embedded in the English text stored in the English text storing means 7, based on the tag and tag information stored in the tag storing means 4, by using the tag embedding means 11, a HTML document with the tag embedded is generated (Step 24), and it is stored in the generated document storing means 13 (Step 25).

[0042]

On the other hand, when it is determined in Step 23 that the kind of text is a translation, the tag is embedded in the

translated text stored in the translated text storing means 9 in a similar manner as in the case of the English text, by determining the position for embedding the tag, based on the corresponding relationship between the English text and the translated text stored in the English-translation corresponding relationship storing means 10, by means of the tag embedding means, a HTML document with the embedded tag in the translated text is created (Step 26), and the process advances to Step 25.
[0043]

Further, when determined in Step 22 that embedding of the tag is not necessary, the process advances to No, and determines a kind of text to be processed (Step 27).
[0044]

When determined in Step 27 that the text to be processed is English, the English text stored in the English text storing means 7 is generated as a HTML document (Step 28), and the process advances to Step 25.
[0045]

On the other hand, when determined in Step 27 that the kind of text to be processed is a translated text, the translated text stored in the translated text storing means 9 is generated as a HTML document (Step 29), and the process advances to Step 25.
[0046]

The above-described operation of Fig. 5 is carried out for all data stored in the English text storing means 7 and the translated text storing means 9.

////////// The End //////////

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平 10 - 134059

(43) 公開日 平成10年(1998)5月22日

(51) Int. Cl. ⁶	識別記号	F I
G 0 6 F 17/28		G 0 6 F 15/38 Q
12/00 5 4 7		12/00 5 4 7 H
// G 0 6 F 13/00 3 5 1		13/00 3 5 1 G

審査請求 未請求 請求項の数 9

O L

(全 8 頁)

(21) 出願番号 特願平8-292331

(22) 出願日 平成8年(1996)11月5日

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 秋山 薫子

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

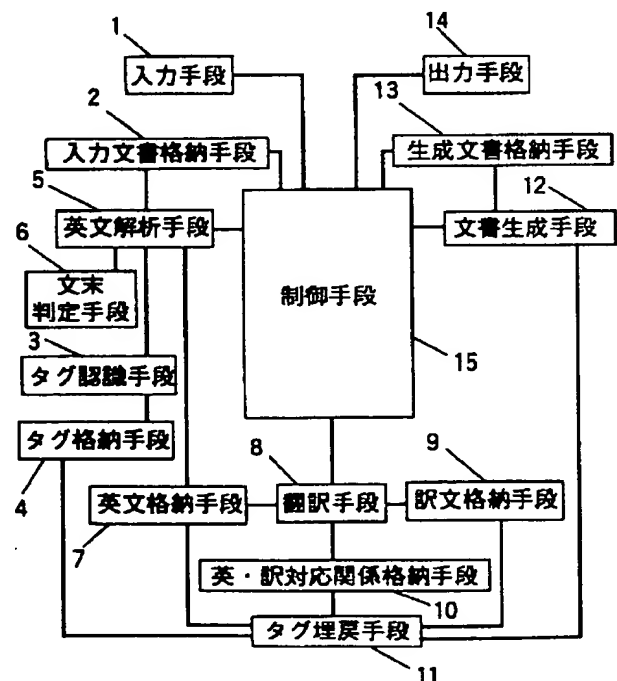
(74) 代理人 弁理士 滝本 智之 (外1名)

(54) 【発明の名称】 HTML 文書処理装置及びHTML 文書処理方法

(57) 【要約】

【課題】 入力されるHTML 文書を解析して翻訳し、原文と訳文とを含むHTML 文書を生成して出力するHTML 文書処理装置を提供すること。

【解決手段】 入力手段 1 から入力されるHTML 文書中のタグをタグ認識手段 3 を用いて認識し、翻訳対象文を取り出す英文解析手段 5 と、認識されたタグ及びタグの位置を格納するタグ格納手段 4 と、取り出された翻訳対象文を格納する英文格納手段 7 と、英文格納手段に格納された翻訳対象文を他の言語に翻訳する翻訳手段 8 と、翻訳結果である訳文を格納する訳文格納手段 9 と、タグ格納手段に格納されたタグ及びタグの位置に基づいて、英文格納手段 7 に格納された翻訳対象文にタグ埋戻手段 11 を用いてタグを埋め戻し、訳文格納手段 9 に格納されている訳文を付加したHTML 文書を生成する文書生成手段 12 と、生成されたHTML 文書を出力する出力手段 14 とを備える。



【特許請求の範囲】

【請求項1】ある言語で書かれたHTML文書を入力する入力手段と、
前記入力手段から入力されるHTML文書を格納する入力文書格納手段と、
前記入力文書格納手段に格納されたHTML文書中のタグを認識して、前記入力文書格納手段に格納されたHTML文書から翻訳対象文を取り出す解析手段と、
前記解析手段により認識されたタグ及びタグの位置を格納するタグ格納手段と、
前記解析手段により取り出された翻訳対象文を格納する原文格納手段と、
前記原文格納手段に格納された翻訳対象文を他の言語に翻訳する翻訳手段と、
前記翻訳手段の翻訳結果である訳文を格納する訳文格納手段と、
前記タグ格納手段に格納されたタグ及びタグの位置に基づいて、前記原文格納手段に格納された翻訳対象文にタグを埋め戻し、前記訳文格納手段に格納されている訳文を付加したHTML文書を生成する文書生成手段と、
前記文書生成手段により生成されたHTML文書を格納する生成文書格納手段と、
前記生成文書格納手段に格納されたHTML文書を出力する出力手段とを備えることを特徴とするHTML文書処理装置。

【請求項2】前記解析手段は、認識したタグに基づいて、翻訳対象文の文末を判断することを特徴とする請求項1記載のHTML文書処理装置。

【請求項3】前記翻訳手段による翻訳の際に、前記原文格納手段に記憶された翻訳対象文と前記訳文格納手段に格納された訳文との対応関係を格納する原・訳対応関係格納手段を備え、
前記文書生成手段は、前記原文格納手段に格納された翻訳対象文にタグを埋め戻すと共に、前記原・訳対応関係格納手段に格納された翻訳対象文と訳文との対応関係に基づいて、翻訳対象文のタグ埋め戻し位置と対応する位置に、前記訳文格納手段に格納された訳文にタグを埋め戻した訳文を付加することを特徴とする請求項1または請求項2記載のHTML文書処理装置。

【請求項4】データを入力する入力装置と、データを記憶する記憶装置と、データを出力する出力装置とを備える装置の制御方法であって、
入力装置から入力されるHTML文書を記憶装置に格納し、
格納されたHTML文書中のタグを認識して、翻訳対象文を取り出し、
認識されたタグ及びタグの位置、並びに取り出された翻訳対象文を記憶装置に格納し、
格納された翻訳対象文を翻訳して、訳文を記憶装置に格納し、

格納されたタグ及びタグの位置に基づいて、格納された翻訳対象文にタグを埋め戻すと共に、格納された訳文を付加したHTML文書を生成し、
生成されたHTML文書を記憶装置に格納し、
格納されたHTML文書を出力装置に出力させるように制御することを特徴とするHTML文書処理方法。

【請求項5】翻訳対象文を取り出す際に、認識したタグに基づいて、翻訳対象文の文末を判断することを特徴とする請求項4記載のHTML文書処理方法。

10 【請求項6】翻訳対象文を翻訳する際に、翻訳対象文と訳文との対応関係を記憶装置に格納し、
格納された訳文を付加する際に、格納された翻訳対象文と訳文との対応関係に基づいて、翻訳対象文のタグ埋め戻し位置と対応する位置に、前記訳文格納手段に格納された訳文にタグを埋め戻した訳文を付加することを特徴とする請求項4または請求項5記載のHTML文書処理方法。

20 【請求項7】データを入力する入力装置と、データを記憶する記憶装置と、データを出力する出力装置とを備える装置を制御するためのプログラムを記憶した記憶媒体であって、

入力装置から入力されるHTML文書を記憶装置に格納し、
格納されたHTML文書中のタグを認識して、翻訳対象文を取り出し、
認識されたタグ及びタグの位置、並びに取り出された翻訳対象文を記憶装置に格納し、
格納された翻訳対象文を翻訳して、訳文を記憶装置に格納し、

30 格納されたタグ及びタグの位置に基づいて、格納された翻訳対象文にタグを埋め戻すと共に、格納された訳文を付加したHTML文書を生成し、
生成されたHTML文書を記憶装置に格納し、
格納されたHTML文書を出力装置に出力させるように制御することを特徴とするプログラムを記憶した記憶媒体。

【請求項8】翻訳対象文を取り出す際に、認識したタグに基づいて、翻訳対象文の文末を判断することを特徴とするプログラムを記憶した請求項7記載の記憶媒体。

40 【請求項9】翻訳対象文を翻訳する際に、翻訳対象文と訳文との対応関係を記憶装置に格納し、
格納された訳文を付加する際に、格納された翻訳対象文と訳文との対応関係に基づいて、翻訳対象文のタグ埋め戻し位置と対応する位置に、前記訳文格納手段に格納された訳文にタグを埋め戻した訳文を付加することを特徴とするプログラムを記憶した請求項6または請求項7記載の記憶媒体。

【発明の詳細な説明】

【0001】

50 【発明の属する技術分野】本発明は、インターネット

(Internet) 上で用いられるHTML (HyperText Markup Language) 言語で書かれた文書を解析し、生成するHTML文書処理装置及びHTML文書処理方法に関するものである。

【0002】

【従来の技術】近年、パーソナルコンピュータ及び通信装置の普及により、インターネットに代表される通信ネットワークを通じて多くの情報を得ることができる。また、インターネットにおける情報システムであるWWW (World Wide Web) では、情報を記述する言語としてHTMLが採用されている。

【0003】

【発明が解決しようとする課題】しかしながら、インターネットは世界規模のネットワークであるため、インターネット上を流れている情報の多くは英文で書かれており、検索ツール（ブラウザ）では、使用者は英文のまま情報に触れなければならなかった。これでは、英語があまり得意でない使用者は、価値ある情報も正確に伝わらない恐れもあり、また、貴重な情報を見逃してしまう恐れもあるといったような問題点があった。

【0004】本発明は、入力されるHTML文書を解析して翻訳し、原文と訳文を含むHTML文書を生成して出力するHTML文書処理装置及びHTML文書処理方法を提供することを目的とする。

【0005】

【課題を解決するための手段】この課題を解決するために本発明は、ある言語で書かれたHTML文書を入力する入力手段と、入力手段から入力されるHTML文書を格納する入力文書格納手段と、入力文書格納手段に格納されたHTML文書中のタグを認識して、入力文書格納手段に格納されたHTML文書から翻訳対象文を取り出す解析手段と、解析手段により認識されたタグ及びタグの位置を格納するタグ格納手段と、解析手段により取り出された翻訳対象文を格納する原文格納手段と、原文格納手段に格納された翻訳対象文を他の言語に翻訳する翻訳手段と、翻訳手段の翻訳結果である訳文を格納する訳文格納手段と、タグ格納手段に格納されたタグ及びタグの位置に基づいて、原文格納手段に格納された翻訳対象文にタグを埋め戻し、訳文格納手段に格納されている訳文を付加したHTML文書を生成する文書生成手段と、文書生成手段により生成されたHTML文書を格納する生成文書格納手段と、生成文書格納手段に格納されたHTML文書を出力する出力手段とを備える。

【0006】これにより、入力されるHTML文書を解析して翻訳し、原文と訳文を含むHTML文書を生成して出力するHTML文書処理装置が得られる。

【0007】

【発明の実施の形態】本発明の請求項1に記載の発明は、ある言語で書かれたHTML文書を入力する入力手段と、入力手段から入力されるHTML文書を格納する

入力文書格納手段と、入力文書格納手段に格納されたHTML文書中のタグを認識して、入力文書格納手段に格納されたHTML文書から翻訳対象文を取り出す解析手段と、解析手段により認識されたタグ及びタグの位置を格納するタグ格納手段と、解析手段により取り出された翻訳対象文を格納する原文格納手段と、原文格納手段に格納された翻訳対象文を他の言語に翻訳する翻訳手段と、翻訳手段の翻訳結果である訳文を格納する訳文格納手段と、タグ格納手段に格納されたタグ及びタグの位置に基づいて、原文格納手段に格納された翻訳対象文にタグを埋め戻し、訳文格納手段に格納されている訳文を付加したHTML文書を生成する文書生成手段と、文書生成手段により生成されたHTML文書を格納する生成文書格納手段と、生成文書格納手段に格納されたHTML文書を出力する出力手段とを備える構成としたことにより、HTML文書中のタグを取り除いて翻訳対象文を翻訳し、翻訳対象文にタグを戻すと共に、訳文を付加した新たなHTML文書を生成することができ、原文と訳文とを同一の表示画面上において見ることができる。

【0008】本発明の請求項2に記載の発明は、請求項1に記載の発明において、解析手段は、認識したタグに基づいて、翻訳対象文の文末を判断する構成としたことにより、より翻訳精度の高い翻訳結果を得ることができる。

【0009】本発明の請求項3に記載の発明は、請求項1または請求項2に記載の発明において、翻訳手段による翻訳の際に、原文格納手段に記憶された翻訳対象文と訳文格納手段に格納された訳文との対応関係を格納する原・訳対応関係格納手段を備え、文書生成手段は、原文格納手段に格納された翻訳対象文にタグを埋め戻すと共に、原・訳対応関係格納手段に格納された翻訳対象文と訳文との対応関係に基づいて、翻訳対象文のタグ埋め戻し位置と対応する位置に、訳文格納手段に格納された訳文にタグを埋め戻した訳文を付加する構成としたことにより、HTML文書中に存在するアンカータグを、訳文中にも設定することができる。

【0010】

【実施例】以下、本発明の一実施例について、図面を参照しながら説明する。

【0011】図1は本発明の一実施例におけるHTML文書処理装置の機能ブロック図である。本実施例では、英語で書かれたHTML文書を解析し、日本語に翻訳したHTML文書を生成することを例として説明する。

【0012】図1において、1はHTML文書の入力や、各種の指示データの入力が行われる入力手段、2は入力手段1から入力されたHTML文書を格納する入力文書格納手段である。

【0013】3は入力文書格納手段2に格納されたHTML文書中に存在するタグを認識するタグ認識手段、4はタグ認識手段3により認識されたタグ及びそのタグが

文書中のどの位置に存在するのかわかる情報（以下「タグ情報」と称する）を格納するタグ格納手段、5はタグ認識手段3を用いてタグを取り除くと共に、タグ格納手段4にタグ情報を格納しながら翻訳対象の英文を取り出す英文解析手段、6はタグ情報格納手段4に格納されたタグ情報に基づいて英文の文末を判断する文末判定手段である。

【0014】7は英文解析手段5により取り出された英文を格納する英文格納手段、8は英文格納手段7に格納された英文を翻訳する翻訳手段、9は翻訳手段8で翻訳された訳文を格納する訳文格納手段、10は英文格納手段7に格納された英文及び訳文格納手段9に格納された訳文に基づいて、英文と訳文との対応関係を記憶する英・訳対応関係格納手段、11はタグ格納手段4に格納されたタグ情報に基づいて、タグを埋め戻すタグ埋戻手段である。

【0015】12はタグ埋戻手段11を用いてタグを埋め戻した英文を生成すると共に、英・訳対応関係格納手段10に格納されている英文と訳文との対応関係に基づいて、英文中に存在するアンカータグを訳文中にも設定した文書を生成する文書生成手段、13は文書生成手段12により生成されたHTML文書を格納する生成文書格納手段である。

【0016】14は生成文書格納手段13に格納されたHTML文書を出力する出力手段、15は前述の各手段のデータの流れを制御する制御手段である。

【0017】図2は本発明の一実施例におけるHTML文書処理装置のハードウェアブロック図である。

【0018】図2において、キーボード21は入力手段1を実現している。ディスプレイ24は出力手段14を実現している。RAM22は入力文書格納手段2、タグ格納手段4、英文格納手段7、訳文格納手段9、英・訳対応関係格納手段10、生成文書格納手段13を実現している。

【0019】CPU25は制御手段15を実現している。ROM23はタグ認識手段3、英文解析手段5、文末判定手段6、翻訳手段8、タグ埋戻手段11、文書生成手段12における各処理を実現するための制御プログラムが格納されている。

【0020】以上のように構成されたHTML文書作成装置について、以下にその動作を説明する。

【0021】図3は本発明の一実施例におけるHTML文書作成装置の動作フローチャートである。

【0022】図3に示すように、入力手段1を介してHTML文書が入力される（ステップ1）。この入力、入力手段1がFD（フロッピーディスク）などの記憶媒体からのロード機能を有しているものであっても良い。

【0023】入力手段1から入力されたHTML文書が入力文書格納手段2にテキスト状態で格納される（ステップ2）。

【0024】英文解析手段5は、入力文書格納手段2に格納された文書データに対して、タグ認識手段3及び文末判定手段6を用いて英文の解析を行い、翻訳対象となる英文を英文格納手段7に格納する。また、タグ認識手段3でタグを認識した際に、タグ情報をタグ格納手段4に格納する（ステップ3）。

【0025】ステップ3における英文解析後、英文格納手段7に格納されている英文を翻訳手段8により翻訳し、翻訳結果である訳文を訳文格納手段9に格納する。また、翻訳手段8が翻訳を行う際に、英文格納手段7に格納されている英文と、訳文格納手段9に格納されている訳文との対応関係を英・訳対応関係格納手段10に格納する（ステップ4）。

【0026】この後、文書生成手段12は、タグ埋戻手段11を用い、タグ格納手段4に格納されているタグの情報に基づいて、英文格納手段7に格納されている英文、及び訳文格納手段9に格納されている訳文にタグを埋め戻してHTML文書を生成し、生成文書格納手段13に格納する（ステップ5）。詳しくは後述する。

【0027】訳文が含まれたHTML文書が生成されると、生成文書格納手段13に格納されたHTML文書を出力手段14に出力する（ステップ6）。

【0028】以上のようにHTML文書処理装置は動作するのであるが、図3のステップ3及びステップ5における各動作を、以下に詳しく説明する。

【0029】図4は本発明の一実施例における英文解析手段のフローチャートであり、図3のステップ3における英文解析の動作を詳しく説明するものである。

【0030】図4において、入力文書格納手段2に格納されている文書データに対して解析処理が終了したのかどうかの判断を行う（ステップ11）。

【0031】ステップ11において、終了したのであればYesに進み、前述した図3のステップ4における翻訳処理に移行する。また、終了していないのであればNoに進み、タグ認識手段3を用いて、文字の種類の判断、つまり、文字が英語であるのかタグであるのかの判断を行う（ステップ12）。

【0032】ステップ12において、文字の種類が英語であると判断すれば、文末判定手段6を用いることにより、その文字が文末の文字であるのかどうかを判断する（ステップ13）。

【0033】ステップ13において、文末であると判断すればYesに進み、英文を取り出し、取り出した英文を翻訳の対象文として英文格納手段7に格納する（ステップ14）。また、ステップ13において、文末でないと判断すればNoに進み、ステップ11に戻る。

【0034】一方、ステップ12において、文字の種類がタグであると判断すれば、タグ情報をタグ格納手段4に格納し（ステップ15）、文末判定手段6を用いて、そのタグが文末のタグであるか否かを判断する（ステ

ブ16)。

【0035】ステップ16において、文末のタグであればYesに進み、ステップ14に移行する。文末のタグでなければNoに進み、ステップ11に戻る。

【0036】以上の図4に示す動作は、入力文書格納手段2に格納されている全てのデータに対して行われる。

【0037】図5は本発明の一実施例における文書生成手段のフローチャートであり、図3のステップ5におけるHTML文書の生成処理の動作を詳しく説明するものである。

【0038】図5に示すように、英文格納手段7及び訳文格納手段9に格納されているデータに対してのHTML文書の生成処理が終了したのかどうかを判断する(ステップ21)。処理が終了した場合は、前述の図3のステップ6における文書の出力処理に移行する。

【0039】ステップ21において、処理が終了していない場合はNoに進み、タグ埋戻手段11を用い、タグ格納手段4に格納されているタグ及びタグ情報に基づいて、タグの埋め戻しが必要であるのかどうかを判断する(ステップ22)。

【0040】ステップ22において、タグの埋め戻しが必要であると判断すれば、Yesに進み、処理の対象となっている文の種類を判断する(ステップ23)。

【0041】ステップ23において、文の種類が英文であると判断すれば、タグ埋戻手段11により、タグ格納手段4に格納されているタグ及びタグ情報に基づいて、英文格納手段7に格納されている英文に対し、タグを埋め戻したHTML文書を生成し(ステップ24)、生成文書格納手段13に格納する(ステップ25)。

【0042】一方、ステップ23において、文の種類が訳文であると判断すれば、英文の場合と同様の要領で、タグ埋戻手段11により、英・訳対応関係格納手段10に格納されている英文と訳文との対応関係に基づいて、タグの埋め戻し位置を決定しながら、訳文格納手段9に格納されている訳文に対してタグを埋め戻したHTML文書を作成し(ステップ26)、ステップ25に進む。

【0043】また、ステップ22において、タグの埋め戻しの必要がないのであればNoに進み、処理の対象となっている文の種類を判断する(ステップ27)。

【0044】ステップ27において、英文であると判断すれば、英文格納手段7に格納されている英文をHTML文書として生成して(ステップ28)、ステップ25に進む。

【0045】一方、ステップ27において、訳文であると判断すれば、訳文格納手段9に格納されている訳文をHTML文書として生成して(ステップ29)、ステップ25に進む。

【0046】以上の図5に示す動作は、英文格納手段7及び訳文格納手段9に格納されている全てのデータに対して行われる。

【0047】以上のように、図3、図4及び図5に示されるフローチャートによって、HTML文書処理装置は動作するのであるが、その動作について、例文を用いて説明する。

【0048】図6は本発明の一実施例における訳文挿入前の表示画面を示す図、図7は本発明の一実施例における訳文挿入前のHTML文書データを示す図、図8は本発明の一実施例における訳文挿入後のHTML文書データを示す図、図9は本発明の一実施例における訳文挿入後の表示画面を示す図である。

【0049】図6のように表示されたHTML文書のテキストデータが図7である。図7に示されるデータに対して解析処理を行うと、タグの認識が行われ、次に示すタグが検出される。

【0050】<FONTSIZE=-1>

このうち、次に示すものがアンカータグと呼ばれるものである。

【0051】

タグが検出されると、次に示すタグ以外の部分が英文であると解析される。

【0052】Coporate Sales:

415/528-2555;

Personal Sales:

415/528-3777

If you have any questions,
please visit Customer Service.

このうち、次の文が翻訳対象として翻訳されることになる。本実施例では英文字を含まない文は翻訳対象から除外している。

【0053】Coporate Sales:

Personal Sales:

If you have any questions,
please visit Customer Service.

これを翻訳した訳文は、次のようになる。

【0054】法人の販売:

パーソナルな販売: もしあなたがあらゆる質問を持っていれば、顧客サービスをどうぞ訪れてください。

【0055】この翻訳結果を反映したHTML文書を生成すると、図8に示す文書のテキストデータが生成され、このテキストデータを表示画面に表示した状態を示したものが図9である。

【0056】なお、HTML文書処理装置において、訳文中にタグを埋め込む際には、タグの挿入位置を決定する必要がある、特に、英文中に存在するアンカータグについては、英・訳対応関係格納手段10を用いて、英文

と訳文との対応関係を考慮する必要がある。

【0057】上記の例で、英文でのアンカータグは、"customer_serice"を囲んでいるので、この英文に対する訳文"顧客サービス"を囲むように、訳文に対してアンカータグを埋め戻す必要がある。このように、英文と訳文との対応関係を考慮して訳文中にタグを入れることにより、英文と訳文とが近接して表示される共に、訳文中にアンカータグを設定することが可能となる。

【0058】なお、本実施例のHTML文書処理装置は、FD、CD-ROM、DVD、HDDなどの記憶媒体に格納されるプログラムをパーソナルコンピュータに装着して格納し、動作させることも可能である。

【0059】

【発明の効果】以上のように本発明によれば、ある言語で書かれたHTML文書を入力する入力手段と、入力手段から入力されるHTML文書を格納する入力文書格納手段と、入力文書格納手段に格納されたHTML文書中のタグを認識して、入力文書格納手段に格納されたHTML文書から翻訳対象文を取り出す解析手段と、解析手段により認識されたタグ及びタグの位置を格納するタグ格納手段と、解析手段により取り出された翻訳対象文を格納する原文格納手段と、原文格納手段に格納された翻訳対象文を他の言語に翻訳する翻訳手段と、翻訳手段の翻訳結果である訳文を格納する訳文格納手段と、タグ格納手段に格納されたタグ及びタグの位置に基づいて、原文格納手段に格納された翻訳対象文にタグを埋め戻し、訳文格納手段に格納されている訳文を付加したHTML文書を生成する文書生成手段と、文書生成手段により生成されたHTML文書を格納する生成文書格納手段と、生成文書格納手段に格納されたHTML文書を出力する出力手段とを備える構成としたことにより、HTML文書中のタグを取り除いて翻訳対象文を翻訳し、翻訳対象文にタグを戻すと共に、訳文を付加した新たなHTML文書を生成することができ、原文と訳文とを同一の表示画面上において見ることができ、利用者の使い勝手が向上する。

【0060】本発明の請求項2に記載の発明は、請求項1に記載の発明において、解析手段は、認識したタグに基づいて、翻訳対象文の文末を判断する構成としたことにより、より翻訳精度の高い翻訳結果を得ることができ、翻訳精度の向上を図ることができる。

【0061】本発明の請求項3に記載の発明は、請求項1または請求項2に記載の発明において、翻訳手段による翻訳の際に、原文格納手段に記憶された翻訳対象文と訳文格納手段に格納された訳文との対応関係を格納する原・訳対応関係格納手段を備え、文書生成手段は、原文格納手段に格納された翻訳対象文にタグを埋め戻すと共に、原・訳対応関係格納手段に格納された翻訳対象文と

訳文との対応関係に基づいて、翻訳対象文のタグ埋め戻し位置と対応する位置に、訳文格納手段に格納された訳文にタグを埋め戻した訳文を付加する構成としたことにより、HTML文書中に存在するアンカータグを、訳文中にも設定することができ、訳文中のアンカータグ部分を指示することによって、他のページを検索することも可能となる。

【図面の簡単な説明】

【図1】本発明の一実施例におけるHTML文書処理装置の機能ブロック図

【図2】本発明の一実施例におけるHTML文書処理装置のハードウェアブロック図

【図3】本発明の一実施例におけるHTML文書作成装置のフローチャート

【図4】本発明の一実施例における英文解析手段のフローチャート

【図5】本発明の一実施例における文書生成手段のフローチャート

【図6】本発明の一実施例における訳文挿入前の表示画面を示す図

【図7】本発明の一実施例における訳文挿入前のHTML文書データを示す図

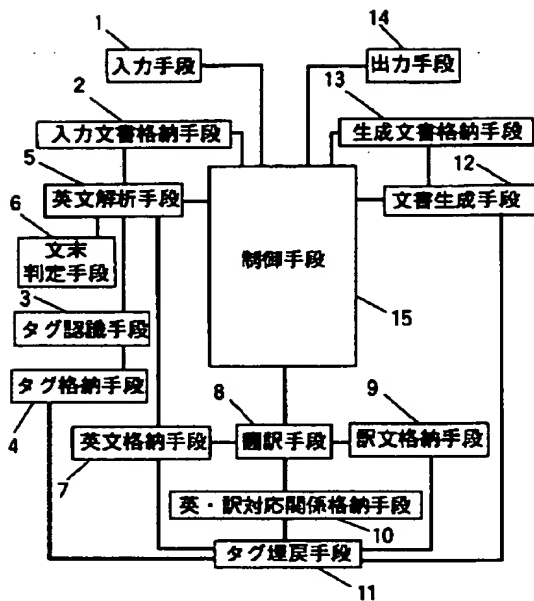
【図8】本発明の一実施例における訳文挿入後のHTML文書データを示す図

【図9】本発明の一実施例における訳文挿入後の表示画面を示す図

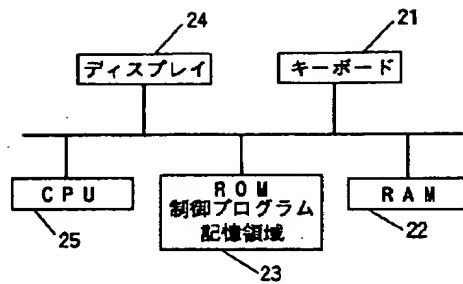
【符号の説明】

- 1 入力手段
- 2 入力文書格納手段
- 3 タグ認識手段
- 4 タグ格納手段
- 5 英文解析手段
- 6 文末判定手段
- 7 英文格納手段
- 8 翻訳手段
- 9 訳文格納手段
- 10 英・訳対応関係格納手段
- 11 タグ埋戻手段
- 12 文書生成手段
- 13 生成文書格納手段
- 14 出力手段
- 15 制御手段
- 21 キーボード
- 22 RAM
- 23 ROM
- 24 ディスプレイ
- 25 CPU

【図1】



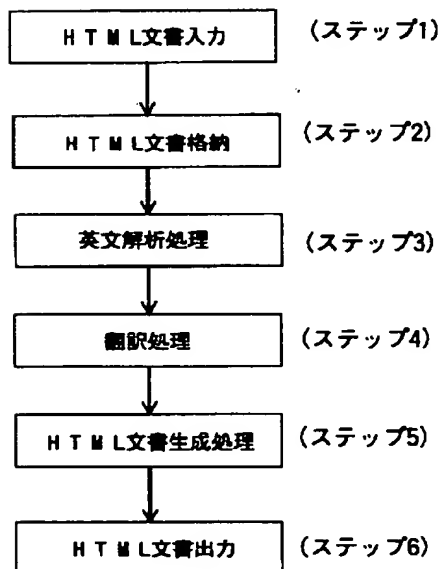
【図2】



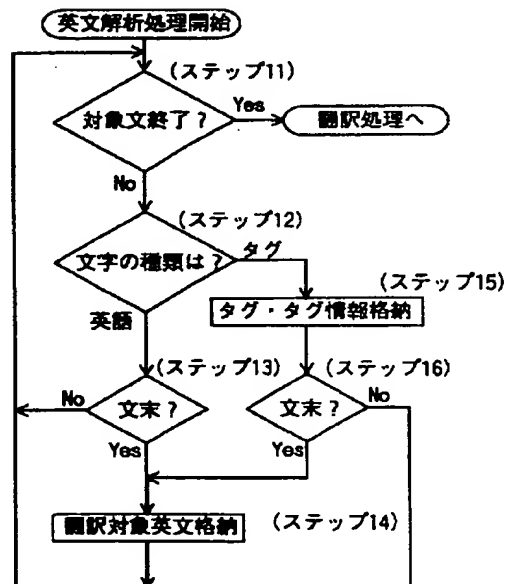
【図6】

Corporate Sales:415/528-2555;Personal Sales:415/528-3777
If you have any questions, please visit Customer Service.

【図3】



【図4】

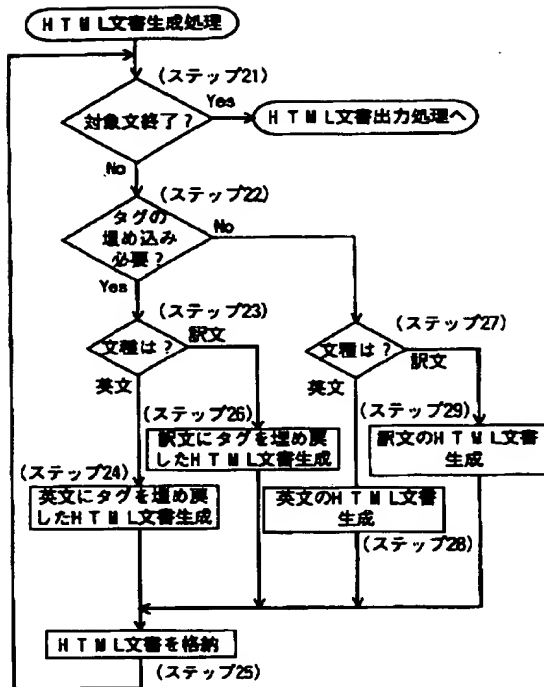


【図7】

Corporate Sales:415/528-2555;Personal Sales:415/528-3777

If you have any questions, please visitCustomer Service.

【図5】



【図9】

Corporate Sales:
法人の販売:
415/528-2555;

Personal Sales:
パーソナルな販売:
415/528-3777

If you have any questions, please visit Customer Service.
もしもあなたがあらゆる質問を持っていれば、顧客サービスをどうぞ訪れてください。

【図8】

Corporate Sales:

法人の販売:

415/528-2555;

Personal Sales:

パーソナルな販売:

415/528-3777

If you have any questions, please visitCustomer Service.

もしもあなたがあらゆる質問を持っていれば、顧客サービスをどうぞ訪れ

てください。
